

STIC-Biotech/ChemLib

164213

From: Mertz, Prema
Sent: Monday, August 29, 2005 2:48 PM
To: STIC-Biotech/ChemLib
Subject: 10/649,857

Please search SEQ ID NO:45 with protein databases.

Thanks

Prema Mertz, Ph.D.
Primary Examiner
Art Unit 1646
4D81 Remsen Bldg Mailbox 4C70
US Patent & Trademark Office
Tel # (571) 272-0876
FAX # (571) 273-0876

CRFE

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked Up: 8/29/05
Date Completed: 8/30/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: 1
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: 207
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2005, 20:32:38 ; Search time 162 Seconds
(without alignments)
162.507 Million cell updates/sec

Title: US-10-649-857-45
Perfect score: 371
Sequence: 1 MKKLCRGAEISLCSFPFL.....PGTTPAPGLFLLLSFWAV 67

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1767149 seqs, 392926209 residues

Total number of hits satisfying chosen parameters: 1767149

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*

10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/2/pubpaa/US09C_NEW_PUB.pep:*

13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*

17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*

18: /cgn2_6/ptodata/2/pubpaa/US10E_NEW_PUB.pep:*

19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*

20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*

21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	371	100.0	67	9	US-09-814-122-45
2	371	100.0	67	15	US-10-649-857-45
3	83.5	22.5	98	16	US-10-425-115-316330
4	74.5	20.1	139	16	US-10-425-115-227490
5	74.5	20.1	200	17	US-10-495-148-12
6	73	19.7	173	15	US-10-424-599-143004
7	71.5	19.3	104	16	US-10-437-963-175436
8	71	19.1	109	16	US-10-425-115-191425
9	71	19.1	672	15	US-10-291-172-587
10	71	19.1	672	15	US-10-291-172-588
11	71	19.1	672	15	US-10-291-172-589

12	71	19.1	672	15	US-10-221-278-587
13	71	19.1	672	15	US-10-221-278-588
14	71	19.1	672	15	US-10-221-278-589
15	70	18.9	87	16	US-10-767-701-62790
16	70	18.9	158	16	US-10-425-115-292945
17	70	18.9	161	16	US-10-425-115-362331
18	69.5	18.7	200	16	US-10-425-115-206766
19	69	18.6	90	16	US-10-437-963-201969
20	69	18.6	241	16	US-10-739-930-6728
21	68.5	18.5	101	16	US-10-425-115-274255
22	68.5	18.5	118	16	US-10-425-115-185747
23	68	18.3	108	16	US-10-425-115-305222
24	68	18.3	118	16	US-10-425-115-231113
25	68	18.3	121	16	US-10-425-115-200190
26	67.5	18.2	115	16	US-10-425-115-288191
27	67.5	18.2	164	15	US-10-424-599-282311
28	67.5	18.2	197	16	US-10-425-115-231743
29	67.5	18.2	270	16	US-10-425-115-191716
30	67.5	18.2	296	15	US-10-425-114-52269
31	67	18.1	145	16	US-10-425-115-223479
32	67	18.1	481	16	US-10-432-934-29
33	66.5	17.9	142	15	US-10-104-047-2674
34	66.5	17.9	145	16	US-10-437-963-184755
35	66.5	17.9	806	15	US-10-161-327-28
36	66	17.8	76	15	US-10-424-599-281813
37	66	17.8	119	16	US-10-437-963-193755
38	66	17.8	517	15	US-10-282-122A-61986
39	66	17.8	525	15	US-10-282-122A-62585
40	66	17.8	525	15	US-10-282-122A-64802
41	66	17.8	712	15	US-10-282-122A-62512
42	66	17.8	720	15	US-10-282-122A-64462
43	65.5	17.7	137	16	US-10-437-963-150511
44	65	17.5	134	15	US-10-424-599-165622
45	65	17.5	186	16	US-10-425-115-320413

ALIGNMENTS

RESULT 1

US-09-814-122-45

; Sequence 45, Application US/09814122

; Patent No. US20020058307A1

; GENERAL INFORMATION:

; APPLICANT: Craig Rosen et al.

; TITLE OF INVENTION: 20 Human Secreted Proteins

; FILE REFERENCE: P2005P1

; CURRENT APPLICATION NUMBER: US/09/814,122

; CURRENT FILING DATE: 2001-03-22

; EARLIER APPLICATION NUMBER: US/09/166,780

; EARLIER FILING DATE: 1998-10-06

; EARLIER APPLICATION NUMBER: PCT/US98/06801

; EARLIER FILING DATE: 1998-04-07

; EARLIER APPLICATION NUMBER: 60/042,726

; EARLIER FILING DATE: 1997-04-08

; EARLIER APPLICATION NUMBER: 60/042,727

; EARLIER FILING DATE: 1997-04-08

; EARLIER APPLICATION NUMBER: 60/042,728

; EARLIER FILING DATE: 1997-04-08

; EARLIER APPLICATION NUMBER: 60/042,754

; EARLIER FILING DATE: 1997-04-08

; EARLIER APPLICATION NUMBER: 60/042,825

; EARLIER FILING DATE: 1997-04-08

; EARLIER APPLICATION NUMBER: 60/048,068

; EARLIER FILING DATE: 1997-05-30

; EARLIER APPLICATION NUMBER: 60/048,070

; EARLIER FILING DATE: 1997-05-30

; EARLIER APPLICATION NUMBER: 60/048,184

; EARLIER FILING DATE: 1997-05-30

; NUMBER OF SEQ ID NOS: 86

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 45

; LENGTH: 67

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-814-122-45

Query Match      100.0%; Score 371; DB 9; Length 67;
Best Local Similarity 100.0%; Pred. No. 6.2e-34;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGKLCRCGAELSCFSPFLLPLHTPVAGRNLGFPESLGVPPFLPHPGGTPRAPGLFL 60
    |||||
Db 1 MGKLCRCGAELSCFSPFLLPLHTPVAGRNLGFPESLGVPPFLPHPGGTPRAPGLFL 60

Qy 61 LFSFWAV 67
    |||||
Db 61 LFSFWAV 67

RESULT 2
US-10-649-857-45
; Sequence 45, Application US/10649857
; Publication No. US20040063128A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen et al.
; TITLE OF INVENTION: 20 Human Secreted Proteins
; FILE REFERENCE: P2005P1
; CURRENT APPLICATION NUMBER: US/10/649,857
; CURRENT FILING DATE: 2003-08-28
; PRIOR FILING DATE: 1998-10-06
; PRIOR FILING DATE: 1998-10-06
; PRIOR FILING DATE: 1998-04-07
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: 60/042,726
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 60/042,727
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 60/042,728
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 60/042,754
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 60/042,825
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: 60/048,068
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,070
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,184
; PRIOR FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-649-857-45

Query Match      100.0%; Score 371; DB 15; Length 67;
Best Local Similarity 100.0%; Pred. No. 6.2e-34;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGKLCRCGAELSCFSPFLLPLHTPVAGRNLGFPESLGVPPFLPHPGGTPRAPGLFL 60
    |||||
Db 1 MGKLCRCGAELSCFSPFLLPLHTPVAGRNLGFPESLGVPPFLPHPGGTPRAPGLFL 60

Qy 61 LFSFWAV 67
    |||||
Db 61 LFSFWAV 67

RESULT 3
US-10-425-115-316330
; Sequence 316330, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 316330
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(98)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_51561C.1.pep
US-10-425-115-316330

Query Match      22.5%; Score 83.5; DB 16; Length 98;
Best Local Similarity 42.0%; Pred. No. 0.14;
Matches 21; Conservative 5; Mismatches 15; Indels 9; Gaps 2;

Qy 16 SFFPFLPLHTPVAGRNLGFP-----ESLGVPPFLPHPGGTPRAPGLF 58
    |||||
Db 51 SFFPNKVPFSPRGLRNSFFLPKTPPNRGRVPP--RNFYSRGRGIF 98

RESULT 4
US-10-425-115-227490
; Sequence 227490, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 227490
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(139)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_139067C.1.pep
US-10-425-115-227490

Query Match      20.1%; Score 74.5; DB 16; Length 139;
Best Local Similarity 37.9%; Pred. No. 2.1;
Matches 22; Conservative 3; Mismatches 20; Indels 13; Gaps 3;

Qy 8 GAELSLCFSPFLLPLHTPVAGRNLGFPESLGVPPFL---PHPGGTPRAPGLFL 62
    |||||
Db 7 GEKEPFCF-FFP-----ENQKLGFPFLDPPFLCFVPPFGKAKSPGPVFLPF 54

RESULT 5
US-10-495-148-12
; Sequence 12, Application US/10495148
; Publication No. US20050107588A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
```

APPLICANT: DUGGAN, Brendan M.
APPLICANT: YANG, Junming
APPLICANT: GIETZEN, Kimberly J.
APPLICANT: LEE, Soo Yeun
APPLICANT: TANG, Y. Tom
APPLICANT: AZIMZAI, Valda
APPLICANT: CHAWLA, Narinder K.
APPLICANT: WARREN, Bridget A.
APPLICANT: BARROSO, Ines
APPLICANT: BECHA, Shanya D.
APPLICANT: YUE, Henry
APPLICANT: LEHR-MASON, Patricia M.
APPLICANT: THANGAVELU, Kavitha
APPLICANT: LEE, Sally
APPLICANT: EMERLING, Brooke M.
APPLICANT: KABLE, Amy E.
APPLICANT: KHARE, Reena
APPLICANT: BAUGHN, Mariah R.
APPLICANT: GANDHI, Ameena R.
APPLICANT: TRAN, Uyen K.
APPLICANT: RICHARDSON, Thomas W.
APPLICANT: MARQUIS, Joseph P.
APPLICANT: LAL, Preeti G.
APPLICANT: FORSYTHE, Ian J.
APPLICANT: LEE, Ernestine A.
APPLICANT: SWARNAKAR, Anita
APPLICANT: KALLICK, Deborah A.
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: CORVAD, Ann E.
APPLICANT: HAFALIA, April J.A.
APPLICANT: ISON, Craig H.
APPLICANT: JIN, Pei
APPLICANT: JIANG, Xin
APPLICANT: JACKSON, Alan A.
APPLICANT: BHATIA, Umesh D.
APPLICANT: BURRILL, John D.
APPLICANT: BLAKE, Julie J.
APPLICANT: HO, Ann
APPLICANT: ZHENG, Wenjin
APPLICANT: GAO, Jing
TITLE OF INVENTION: RECEPTORS AND MEMBRANE-ASSOCIATED PROTEINS
FILE REFERENCE: PF-1279 PCT
CURRENT APPLICATION NUMBER: US/10/495,148
CURRENT FILING DATE: 2004-05-10
PRIOR APPLICATION NUMBER: US 60/333,097
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,274
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/340,542
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/342,166
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: US 60/347,580
PRIOR FILING DATE: 2002-01-11
PRIOR APPLICATION NUMBER: US 60/348,687
PRIOR FILING DATE: 2002-01-14
NUMBER OF SEQ ID NOS: 94
SOFTWARE: PERL Program
SEQ ID NO 12
LENGTH: 200
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 7762537CD1
US-10-495-148-12

Query Match 20.1%; Score 74.5; DB 17; Length 200;
Best Local Similarity 41.2%; Pred. No. 3.2;
Matches 21; Conservative 6; Mismatches 17; Indels 7; Gaps 3;
QY 10 ELSLCFSFPLLLPLHTPVAGRNIGFPPSLGVPPFLPHPGGTPRAPGLFLL 60

Db 65 QLQPCATAYVP-VYPVGTYPYAG---GTPCGTGVSTLPPPP---PQCPGLALL 108
RESULT 6
US-10-424-599-143004
; Sequence 143004, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 143004
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100146C.1.pap
US-10-424-599-143004

Query Match 19.7%; Score 73; DB 15; Length 173;
Best Local Similarity 40.0%; Pred. No. 4;
Matches 20; Conservative 4; Mismatches 20; Indels 6; Gaps 2;
QY 16 SFFPLLPLHTPVAGRNIGFPPSLGVPPFLPHPGGTPRAPGLFLLFSFW 65
Db 110 SMAPVSRPVGTPLARRPSSFP--GGPP---PWGGHPVPVLYLPSFLW 153

RESULT 7
US-10-437-963-175436
; Sequence 175436, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 175436
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_73281C.1.pap
US-10-437-963-175436

Query Match 19.3%; Score 71.5; DB 16; Length 104;
Best Local Similarity 42.1%; Pred. No. 3.4;
Matches 16; Conservative 6; Mismatches 15; Indels 1; Gaps 1;
QY 30 GRNLGFPESLGVPPFLPHPGGTPRAPGLFLLF-SFWA 66
Db 57 GKVPFPGKGLPEPPPPPGAPPSPSWKISPSFWA 94
RESULT 8

```

US-10-425-115-191425
; Sequence 191425, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 191425
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_10615C.1.pep
US-10-425-115-191425

Query Match          19.1%; Score 71; DB 16; Length 109;
Best Local Similarity 28.4%; Pred. NO. 4.1;
Matches 19; Conservative 13; Mismatches 29; Indels 6; Gaps 2;

QY      5  LCRGAEI-----SICFSFPELLPLHTPVAGNLCGFESLGVPPPLP--GGTPRAGLGF 58
          |||      |||      |||      |||      |||      |||      |||      |||      |||
DB      38  LCMGVGFPMSPPTPMTSPSPISPSRSHRGWQPQAAHQHPQYPLLLFPTVRSPGVV 97
          |||      |||      |||      |||      |||      |||      |||      |||      |||
QY      59  LLLLSFW 65
          |||      |
DB      98  VVIQHVW 104

RESULT 9
US-10-291-172-587
; Sequence 587, Application US/10291172
; Publication No. US20030228584A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20030228584A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/291,172
; CURRENT FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 587
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(672)
; OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-291-172-587

Query Match          19.1%; Score 71; DB 15; Length 672;
Best Local Similarity 40.7%; Pred. NO. 29;
Matches 22; Conservative 1; Mismatches 17; Indels 14; Gaps 3;

QY      19  PLLPLPLHTPVAGNLCGF--PESLG-----VPPFLPHPGGT-----PRAPGLGF 58

```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(672)
; OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-291-172-589

Query Match      19.1%; Score 71; DB 15; Length 672;
Best Local Similarity 40.7%; Pred. No. 29;
Matches 22; Conservative 1; Mismatches 17; Indels 14; Gaps 3;

QY 19 PLLPLHTPVAGRNIGF--PESLG-----VPPFLPHPGGT----PRAPGLF 58
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 589 PLQLPFPHPPLGLACGFLPSSVGGRLCGGFVPCFLGAQGAWFHPRPPTAF 642

RESULT 12
US-10-221-278-587
; Sequence 587, Application US/10221278
; Publication No. US20040034208A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/221,278
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 587
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(672)
; OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-221-278-587

Query Match      19.1%; Score 71; DB 15; Length 672;
Best Local Similarity 40.7%; Pred. No. 29;
Matches 22; Conservative 1; Mismatches 17; Indels 14; Gaps 3;

QY 19 PLLPLHTPVAGRNIGF--PESLG-----VPPFLPHPGGT----PRAPGLF 58
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 589 PLQLPFPHPPLGLACGFLPSSVGGRLCGGFVPCFLGAQGAWFHPRPPTAF 642

RESULT 13
US-10-221-278-588
; Sequence 588, Application US/10221278
; Publication No. US20040034208A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/221,278
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 587
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(672)
; OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-221-278-587

Query Match      19.1%; Score 71; DB 15; Length 672;
Best Local Similarity 40.7%; Pred. No. 29;
Matches 22; Conservative 1; Mismatches 17; Indels 14; Gaps 3;

QY 19 PLLPLHTPVAGRNIGF--PESLG-----VPPFLPHPGGT----PRAPGLF 58
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 589 PLQLPFPHPPLGLACGFLPSSVGGRLCGGFVPCFLGAQGAWFHPRPPTAF 642

RESULT 14
US-10-221-278-589
; Sequence 589, Application US/10221278
; Publication No. US20040034208A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: No. US20040034208A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-045
; CURRENT APPLICATION NUMBER: US/10/221,278
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: 09/693,267
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/665,363
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/616,847
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 09/596,193
; PRIOR FILING DATE: 2000-06-17
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 752
; SEQ ID NO 589
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(672)
; OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-221-278-589

Query Match      19.1%; Score 71; DB 15; Length 672;
Best Local Similarity 40.7%; Pred. No. 29;
Matches 22; Conservative 1; Mismatches 17; Indels 14; Gaps 3;

QY 19 PLLPLHTPVAGRNIGF--PESLG-----VPPFLPHPGGT----PRAPGLF 58
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 589 PLQLPFPHPPLGLACGFLPSSVGGRLCGGFVPCFLGAQGAWFHPRPPTAF 642

RESULT 15
US-10-767-701-62790
; Sequence 62790, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
```

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 62790
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 18065573.pep
US-10-767-701-62790

Query Match 18.9%; Score 70; DB 16; Length 87;
Best Local Similarity 51.4%; Pred. No. 4.1;
Matches 19; Conservative 1; Mismatches 5; Indels 12; Gaps 3;
Qy 19 PLLLP--HTPVAGNLTGPPESLGVPPPLPHPGGTPR 53
| | | | | | | | | | | | | | | | | | | | | |
Db 16 PLLLP--HTPVAGNLTGPPESLGVPPPLPHPGGTPR 42
| | | | | | | | | | | | | | | | | | | | | |

Search completed: August 29, 2005, 20:48:13
Job time : 163 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2005, 20:45:34 ; Search time 168 Seconds
(without alignments)
204.222 Million cell updates/sec

Title: us-10-649-857-45

Perfect score: 371

Sequence: i MGKLCRGAEUSLCSFPPL.....PGGTPRAPGLFLLLSFMAV 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74.5	20.1	306	1 NG5 HUMAN	Q99946 homo sapien
2	73	19.7	200	2 Q6ZR36	Q6Zr36 homo sapien
3	72.5	19.5	306	1 NG5 MOUSE	O35449 mus musculus
4	71.5	19.3	297	2 O17662	O17662 caenorhabdi
5	71.5	19.3	1112	2 Q8CGT8	Q8CGT8 mus musculus
6	70.5	19.0	216	2 Q9HEK5	Q9HEK5 neurospora
7	70.5	19.0	313	2 Q976R7	Q976R7 sulfolobus
8	70	18.9	142	2 P90815	P90815 caenorhabdi
9	69	18.6	123	2 Q9BP05	Q9BP05 homo sapien
10	69	18.6	241	2 Q9FIS4	Q9FIS4 arabidopsis
11	68.5	18.5	1802	2 Q17163	Q17163 brugia mala
12	68	18.3	85	2 Q8K4W4	Q8K4W4 mus musculus
13	68	18.3	123	2 Q8K4W5	Q8K4W5 mus musculus
14	67	18.1	481	2 Q07787	Q07787 mycobacteri
15	67	18.1	481	2 Q7U1P5	Q7U1P5 mycobacteri
16	66.5	17.9	145	2 Q94dt5	Q94dt5 oryza sativ
17	66.5	17.9	300	2 Q6MG82	Q6MG82 rattus norv
18	66.5	17.9	420	1 BD12 EBV	P03225 Epstein-bar
19	66.5	17.9	420	2 Q777C4	Q777C4 human herpe
20	66.5	17.9	703	2 Q87MM3	Q87mm3 vibrio para
21	66.5	17.9	718	2 Q6SK49	Q6SK49 arthrobacte
22	66	17.8	123	2 Q9WVJ1	Q9WVJ1 mus musculus
23	66	17.8	124	2 Q9Z2F1	Q9Z2F1 rattus norv
24	66	17.8	381	2 Q6PAL1	Q6PAL1 mus musculus
25	66	17.8	411	2 Q6WJ00	Q6Wj00 gallus gall
26	66	17.8	486	2 Q7RWM4	Q7Rwm4 neurospora
27	66	17.8	517	2 Q73VNO	Q73vno mycobacteri
28	66	17.8	525	1 SR54 MYCBO	P66845 mycobacteri
29	66	17.8	525	1 SR54 MYCTU	P66844 mycobacteri
30	66	17.8	567	1 F2D7 CHICK	O57329 gallus gall
31	66	17.8	589	2 Q63094	Q63094 rattus norv

32	66	17.8	720	2	Q7D952	Q7D952 mycobacteri
33	66	17.8	720	2	O53872	O53872 mycobacteri
34	66	17.8	720	2	Q7U133	Q7U133 mycobacteri
35	66	17.8	1859	2	Q6V1N9	Q6V1n9 streptomyce
36	65.5	17.7	244	2	Q6N6E5	Q6N6e5 rhodospseudo
37	65.5	17.7	257	2	Q9RV14	Q9RV14 deinococcus
38	65.5	17.7	282	2	Q99CY7	Q99CY7 bovine herp
39	65.5	17.7	367	2	Q7S5E3	Q7S5e3 neurospora
40	65.5	17.7	437	2	Q7NYK3	Q7nyk3 chromobacte
41	65.5	17.7	705	2	Q7MIS5	Q7mis5 vibrio vuln
42	65.5	17.7	705	2	Q8DB47	Q8db47 vibrio vuln
43	65	17.5	383	2	O62543	O62543 geodia cydo
44	65	17.5	424	2	Q6VMK7	Q6VMK7 rattus ratt
45	65	17.5	452	2	Q6ZUN0	Q6zun0 homo sapien

ALIGNMENTS

RESULT 1
NG5_HUMAN STANDARD; PRT; 306 AA.
AC Q99946; Q96DM3; Q96NQ8.
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Protein NG5.
GN Name=C6orf31; Synonyms=NG5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Rowen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,
RA Banta A., Spies T., Hood L.;
RT "Sequence determination of 300 kilobases of the human class III MHC
RT locus".
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cerebellum;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsura N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoa K., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiyama S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Itoh T., Shigeta K., Senba T.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";

[illegible]

```
FT DOMAIN 41 49 Poly-His.
FT DOMAIN 121 127 Poly-Pro.
FT DOMAIN 128 131 Poly-Ala.
FT DOMAIN 132 139 Poly-Pro.
FT DOMAIN 203 206 Poly-Pro.
SQ SEQUENCE 306 AA; 31389 MW; D8875395737F3E6B CRC64;

Query Match 19.5%; Score 72.5; DB 1; Length 306;
Best Local Similarity 41.2%; Pred. No. 9.7;
Matches 21; Conservative 6; Mismatches 17; Indels 7; Gaps 3;

QY 10 ELSLCFSPPLLLPLHTPVAGNLFPPSLGVPPFLPHPGGTTPRAPGLFLL 60
Db 171 QLOPCTAYVP-VYPGTPYAG---GTPGPGVTSTLPPP---PQPGGLALL 214

RESULT 4
O17662 PRELIMINARY; PRT; 297 AA.
ID O17662
AC O17662;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DE 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein C41G6.9.
GN ORFNames=C41G6.9;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Cummings P.N.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; 281047; CAB02834.1; -.
DR PIR; T19888; T19888.
DR WormBase; WBGene0005381; C41G6.9.
DR WormPep; C41G6.9; CE15706.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004930; F:G-protein coupled receptor activity; IEA.
DR InterPro; IPR003003; 7TM_chemrecept2.
DR Pfam; PF01604; 7tm 5; 1.
KW Hypothetical protein.
SQ SEQUENCE 297 AA; 33531 MW; 14F8EC5873F534CA CRC64;

Query Match 19.3%; Score 71.5; DB 2; Length 297;
Best Local Similarity 51.2%; Pred. No. 12;
Matches 22; Conservative 3; Mismatches 11; Indels 7; Gaps 3;

QY 4 CL-CRGAELSLCFSPF---PLLLPLHTPVAGNLFPPSLGV 42
Db 61 CLHCVGAFVLDLFFSFPAIPALNLPY---AGVFLGFSRVLGVP 100

RESULT 5
Q8CGT8 PRELIMINARY; PRT; 1112 AA.
ID Q8CGT8
AC Q8CGT8;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Argonaute 4 protein.
GN Name=Eif2c4; Synonyms=Ago4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22301763; PubMed=12414724; DOI=10.1101/gad.1026102;
RA Carmell M.A., Xuan Z., Zhang M.Q., Hannon G.J.;
RT "The Argonaute family: tentacles that reach into RNAi, developmental
RT control, stem cell maintenance, and tumorigenesis.";
RL Genes Dev. 16:2733-2742(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Adams M., Mural R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR MGD; AY135690; AAN75581.1; -.
DR MGD; MGI:1924100; Bif2c4.
DR InterPro; IPR003100; PAZ.
DR InterPro; IPR003165; Piwi.
DR Pfam; PF02170; PAZ; 1.
DR Pfam; PF02171; Piwi; 1.
DR PROSITE; PS50822; PIWI; 1.
SQ SEQUENCE 1112 AA; 124214 MW; AADB6F618BF696B80 CRC64;

Query Match 19.3%; Score 71.5; DB 2; Length 1112;
Best Local Similarity 37.1%; Pred. No. 45;
Matches 23; Conservative 4; Mismatches 20; Indels 15; Gaps 3;

QY 2 GKCGAELSLCFSPFLLPLHT-PVAGNLFPPSLGVPPFLPHPGGTTPRAP 55
Db 116 GCGPALHGGSGELLCSLVLPLWLVAVIKIVTKNLGPPASLFP-----PRRP 166

QY 56 GL 57
Db 167 GL 168

RESULT 6
Q9HEK5 PRELIMINARY; PRT; 216 AA.
ID Q9HEK5
AC Q9HEK5;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Related to UI SMALL NUCLEAR RIBONUCLEOPROTEIN C.
GN Name=12F11.220;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hobeisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL451017; CAC18231.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030529; C:ribonucleoprotein complex; IEA.
DR GO; GO:0019013; C:viral nucleocapsid; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR InterPro; IPR000690; Znf_matin.
DR InterPro; IPR003604; Znf_U1.
DR Pfam; PF06220; zf-U1; 1.
DR SMART; SM00451; Znf_U1; 1.
DR PROSITE; PS50171; ZF_MATRIN; 1.
KW Nucleocapsid; Ribonucleoprotein.
SQ SEQUENCE 216 AA; 21527 MW; D029E2DA62575FEA CRC64;

Query Match 19.0%; Score 70.5; DB 2; Length 216;
Best Local Similarity 47.5%; Pred. No. 11;
```


Matches 16; Conservative 4; Mismatches 19; Indels 4; Gaps 1;

QY 27 PVAGNRLGPPESLGVPPPLPHGPGTPRAPGL---FLLLFSW 65
 Db 56 PPMGPPIGLPARGTPIGNPPGMPPPPGIRGEALLALVCW 98

RESULT 10
 Q9FIS4 ID Q9FIS4 PRELIMINARY; PRT; 241 AA.
 AC Q9FIS4;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DE Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MTG10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99087489; PubMed=9872454;
 RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
 RT Structure features of the regions of 1,013,767 bp covered by sixteen
 RT physically assigned P1 and TAC clones.";
 RL DNA Res. 5:297-308(1998).
 DR EMBL; AB016880; BAB10176.1; -- F8920D05D4E6DF6F CRC64;
 SQ SEQUENCE 241 AA; 26439 MW; F8920D05D4E6DF6F CRC64;

Query Match 18.6%; Score 69; DB 2; Length 241;
 Best Local Similarity 43.2%; Pred. No. 18;
 Matches 16; Conservative 6; Mismatches 11; Indels 4; Gaps 1;

QY 9 AELSLCFSFFPLLLPLHTPVAGRNLPESLGVPPPL 45
 Db 97 SELSVCFNVATLYPSLTSTQTKPLGFP---LPFPL 129

RESULT 11
 Q17163 ID Q17163 PRELIMINARY; PRT; 1802 AA.
 AC Q17163;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE A2 (IV) basement membrane collagen.
 OS Brugia malayi (Filarial nematode worm).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
 OC Onchocercidae; Brugia.
 OX NCBI_TaxID=6279;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=aperiodic;
 RX MEDLINE=95364846; PubMed=76377709; DOI=10.1016/0166-6851(95)00024-U;
 RA Caulagi V.R., Rajan T.V.;
 RT "The structural organization of an alpha 2 (type IV) basement membrane
 RT collagen gene from the filarial nematode Brugia malayi.";
 RL Mol. Biochem. Parasitol. 70:227-229(1995).
 DR EMBL; U07224; AAC46611.1; --
 DR HSP; P08572; IL11.
 DR GO; GO:0005581; C:collagen; IEA.
 DR GO; GO:0005737; C:cytoplasm; IEA.
 DR GO; GO:0005201; F:extracellular matrix structural constituent; IEA.
 DR GO; GO:0006817; P:phosphate transport; IEA.
 DR InterPro; IPR008160; Collagen.
 DR InterPro; IPR001442; Procollagen4_C.
 DR Pfam; PF01413; C4; 2.
 DR Pfam; PF01391; Collagen; 26.
 DR SMART; SM00111; C4; 2.
 KW Collagen.

SQ SEQUENCE 1802 AA; 172402 MW; 595F16554CBE2D24 CRC64;

Query Match 18.5%; Score 68.5; DB 2; Length 1802;
 Best Local Similarity 51.2%; Pred. No. 1.5e+02;
 Matches 22; Conservative 0; Mismatches 16; Indels 5; Gaps 3;

QY 18 FPLL--LPLHTPVAGRNLPFP--BSLGVPPPLPHGPGTPRAPG 56
 Db 1522 FFGLPGLPGEKGAAGLP--GFGVGEVTPGPGPLPGPGPPGAPG 1563

RESULT 12
 Q8K4W4 ID Q8K4W4 PRELIMINARY; PRT; 85 AA.
 AC Q8K4W4;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Calmodulin-dependent protein kinase II beta 3 isoform insert
 DE (Fragment).
 GN Name=Cank2b;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RX PubMed=12110572; DOI=10.1093/emboj/cdf360;
 RA Bayer K.U., De Koning P., Schulman H.;
 RT "Alternative splicing modulates the frequency-dependent response of
 RT CaMKII to Ca²⁺ oscillations.";
 RL EMBO J. 21:3590-3597(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RA Bayer K.-U.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF416337; AAM77845.1; --
 DR EMBL; AF416336; AAM77845.1; JOINED.
 DR GO; GO:0016301; F:kinase activity; IEA.
 KW Kinase.
 FT NON_TER 1 1
 FT NON_TER 85 85
 SQ SEQUENCE 85 AA; 8444 MW; BFIAB9082A0E5E36 CRC64;

Query Match 18.3%; Score 68; DB 2; Length 85;
 Best Local Similarity 33.9%; Pred. No. 8.4;
 Matches 21; Conservative 3; Mismatches 18; Indels 20; Gaps 2;

QY 4 CLCGAELSLCFSFFPLLLPLHTP-----VAGRNLPESLGVPPPLPHGPGTPRA 54
 Db 30 CLSPG-----LLGLPTSPRISDILNSVSGSGTPEAEGLPFGVPPCPSPTL 78

RESULT 13
 Q8K4W5 ID Q8K4W5 PRELIMINARY; PRT; 123 AA.
 AC Q8K4W5;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Calmodulin-dependent protein kinase II beta M isoform insert
 DE (Fragment).
 GN Name=Cank2b;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX PubMed:12110572; DOI=10.1093/emboj/cdf360;
RT Bayer K.U., De Koninck P., Schulman H.;
RA "Alternative splicing modulates the frequency-dependent response of
RT CamKII to Ca2+ oscillations.";
RL EMOB J. 21:3590-3597(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Bayer K.-U.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF416337; AAM77844.1; -.
DR EMBL; AF416336; AAM77844.1; JOINED.
DR GO; GO:0016301; P:kinase activity; IEA.
KW Kinase.
FT NON_TER 1 1
FT NON_TER 123 123
SQ SEQUENCE 123 AA; 12304 MW; C45A4D20531B9347 CRC64;
Query Match 18.3%; Score 68; DB 2; Length 123;
Best Local Similarity 33.9%; Pred. No. 12;
Matches 21; Conservative 3; Mismatches 18; Indels 20; Gaps 2;

Qy 4 CLCRAELSCFPPFLLLPLHTP-----VAGRNLPFESLGVPFPLPHPGGTTPRA 54
Db 68 CLSPG-----LLGLPTSPRISDIILSVSGSGTPEAGLPPVGVPPCPSPPTL 116

Qy 55 PG 56
Db 117 PG 118

RESULT 14
O07787
ID O07787 PRELIMINARY; PRT; 481 AA.
AC O07787;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MCE-FAMILY PROTEIN MCE2C.
GN Name=mce2C; OrderedLocusNames=Rv0591;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Blosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
RA Tekaita P., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd L.D., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.B., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
DR EMBL; BX842574; CAB09961.1; -.
DR PIR; B70908; B70908.
DR TubercuList; Rv0591; -.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR005693; Mce.
DR InterPro; IPR008360; Mce1C.
DR InterPro; IPR003399; Mce_related.
DR Pfam; PF02470; MCE; 1.
DR PRINTS; PR01782; MCEVIRFACTOR.
DR TIGRFAMs; TIGR00996; Mtu_fam_mce; 1.
KW Complete proteome.
```

```

SQ SEQUENCE 481 AA; 50764 MW; 726B8265DDCA6A29 CRC64;
Query Match 18.1%; Score 67; DB 2; Length 481;
Best Local Similarity 40.5%; Pred. No. 60;
Matches 15; Conservative 3; Mismatches 13; Indels 6; Gaps 1;

Qy 27 PVAGRNLG-----FPESLGVPFPLPHPGGTTPRAPGL 57
Db 400 FCAGATVGPFGPDFFAPLDVQPSPPNPDGPPPTPGI 436

RESULT 15
O7UIP5
ID O7UIP5 PRELIMINARY; PRT; 481 AA.
AC O7UIP5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE MCE-FAMILY PROTEIN MCE2C.
GN Name=mce2C; OrderedLocusNames=Mb0606;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eigmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248336; CAD93468.1; -.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR005693; Mce.
DR InterPro; IPR003399; Mce_related.
DR Pfam; PF02470; MCE; 1.
DR TIGRFAMs; TIGR00996; Mtu_fam_mce; 1.
KW Complete proteome.
SQ SEQUENCE 481 AA; 50807 MW; 4CC18C359266B1B8 CRC64;
Query Match 18.1%; Score 67; DB 2; Length 481;
Best Local Similarity 40.5%; Pred. No. 60;
Matches 15; Conservative 3; Mismatches 13; Indels 6; Gaps 1;

Qy 27 PVAGRNLG-----FPESLGVPFPLPHPGGTTPRAPGL 57
Db 400 FCAGATVGPFGPDFFAPLDVQPSPPNPDGPPPTPGI 436

Search completed: August 29, 2005, 20:56:02
Job time : 171 secs
```

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	64.5	17.4	426	4	US-09-489-039A-11995	Sequence 11995, A
2	63.5	17.1	751	4	US-09-543-681A-4181	Sequence 4181, Ap
3	63.5	17.1	847	4	US-09-949-016-6222	Sequence 6222, Ap
4	63.5	17.1	885	4	US-09-949-016-7789	Sequence 7789, Ap
5	63	17.1	265	1	US-08-155-171B-41	Sequence 41, Appl
6	63	17.0	265	2	US-08-435-998-41	Sequence 41, Appl
7	63	17.0	277	4	US-09-431-887-34	Sequence 34, Appl
8	63	17.0	582	4	US-09-252-991A-36182	Sequence 26182, A
9	62.5	16.8	428	4	US-09-252-991A-19452	Sequence 19452, A
10	62	16.7	248	4	US-09-949-016-10554	Sequence 10554, A
11	62	16.7	249	4	US-09-917-254-96	Sequence 96, Appl
12	62	16.7	515	4	US-09-252-991A-38127	Sequence 38127, A
13	61.5	16.6	169	4	US-09-489-039A-8371	Sequence 8371, Ap
14	61.5	16.6	305	4	US-09-252-991A-18762	Sequence 18762, A
15	61.5	16.6	426	4	US-09-270-767-37797	Sequence 37797, A
16	61.5	16.6	426	4	US-09-270-767-53014	Sequence 53014, A
17	61.5	16.6	664	4	US-09-252-991A-29360	Sequence 29360, A
18	61.5	16.6	811	4	US-09-819-989-2	Sequence 2, Appli
19	61.5	16.6	811	4	US-10-273-992-2	Sequence 2, Appli
20	61.5	16.6	811	4	US-10-681-223-2	Sequence 2, Appli
21	61.5	16.6	1572	4	US-09-902-540-13652	Sequence 12652, A
22	61	16.4	240	4	US-09-538-092-929	Sequence 929, App
23	61	16.4	240	4	US-09-538-092-930	Sequence 930, App
24	61	16.4	240	4	US-09-949-016-6314	Sequence 6314, Ap
25	61	16.4	272	4	US-09-949-016-11309	Sequence 11309, A
26	61	16.4	689	4	US-09-252-991A-31332	Sequence 31332, A
27	60.5	16.3	406	4	US-09-270-767-44410	Sequence 44410, A

```

RESULT 2
US-09-543-681A-4181
; Sequence 4181, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4181
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4181
Query Match      17.1%   Score 63.5   DB 4:   Length 751:

```



```
RESULT 10
US-09-949-016-10554
; Sequence 1054, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10554
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10554

Query Match      16.7%; Score 62; DB 4; Length 248;
Best Local Similarity 40.0%; Pred. No. 10;
Matches 14; Conservative 3; Mismatches 18; Indels 0; Gaps 0;

Qy 27 PVAGNIGFPESLGVPPFLPHPGTTPRAPGLFLL 61
Db 214 PPMGPPMGIPGRTGMPGMPGMPGMPGMRGLL 248

RESULT 11
US-09-917-254-96
; Sequence 96, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; APPLICANT: Baak, Jan
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224(JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 96
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-917-254-96

Query Match      16.7%; Score 62; DB 4; Length 249;
Best Local Similarity 40.0%; Pred. No. 10;
Matches 14; Conservative 3; Mismatches 18; Indels 0; Gaps 0;

Qy 27 PVAGNIGFPESLGVPPFLPHPGTTPRAPGLFLL 61
Db 215 PPMGPPMGIPGRTGMPGMPGMPGMPGMRGLL 249

RESULT 12
US-09-252-991A-28127
; Sequence 28127, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

RESULT 13
US-09-489-039A-8371
; Sequence 8371, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8371
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (10),(11),(12),(13),(14),(15),(16),(17),(18),(19),(20),(21),(22)
; LOCATION: (23),(24),(25),(26),(27),(28),(29),(30),(31),(32),(33),(34),(35)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-489-039A-8371

Query Match      16.8%; Score 61.5; DB 4; Length 169;
Best Local Similarity 40.0%; Pred. No. 7;
Matches 18; Conservative 2; Mismatches 22; Indels 3; Gaps 1;

Qy 17 FFPLLPLHTPVAGRNIGFPESLGVPPFLPHPGTTPRAPGLFLL 61
Db 100 FFSLSLFFSPSPSPSLFSPSS---PPLLSPLSLSPFLPLLLL 141

RESULT 14
US-09-252-991A-18762
; Sequence 18762, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
```

```
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18762
; LENGTH: 305
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18762

Query Match      16.6%; Score 61.5; DB 4; Length 305;
Best Local Similarity 44.8%; Pred. No. 15;
Matches 13; Conservative 3; Mismatches 12; Indels 1; Gaps 1;

QY 27 PVAGNLCFPESLGVPPPLPHPGGTTPRAP 55
Db 63 PIAGRRADDPG-GIPFRHHPRRPRSP 90

RESULT 15
US-09-270-767-37797
; Sequence 37797, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37797
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-37797

Query Match      16.6%; Score 61.5; DB 4; Length 426;
Best Local Similarity 32.2%; Pred. No. 22;
Matches 29; Conservative 3; Mismatches 23; Indels 35; Gaps 5;

QY 4 CLC-----RGAELSLCFSEFP---LLPLHTPVAGNLCFPESLGVPPPLPH----- 47
Db 87 CLCPPLTLRLPRCVLPPLSPSFLPAPLRLTLTLLSP-CPRDLSEFP-SPSPVPSVPSCHSV 144
QY 48 -----PGGTTPRAPGLFL 60
Db 145 STDSPSPDCGAYLLVHTPRETPVLELFL 174
```

Search completed: August 29, 2005, 20:37:03
Job time : 44 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2005, 20:37:09 ; Search time 39 Seconds
(without alignments)
165.295 Million cell updates/sec

Title: us-10-649-857-45
Perfect score: 371
Sequence: 1 MGKCLCRGABLSLCSFPPL.....PGGTPRAPGLFLLLFSFWAV 67

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72.5	19.5	306	T09067	extensin-like prot
2	71.5	19.3	297	T19888	hypothetical prote
3	70	18.9	142	T23485	hypothetical prote
4	69.5	18.7	327	JCS319	macrolide-lincosam
5	67	18.1	481	B70908	hypothetical prote
6	66.5	17.9	420	Q0BE44	BDLP2 protein - hu
7	66	17.8	525	D70747	probable ffh prote
8	66	17.8	589	S68470	Ca2+/calmodulin-de
9	66	17.8	720	T09815	probable fadB prot
10	65.5	17.7	257	E75447	hypothetical prote
11	64.5	17.4	102	E72501	hypothetical prote
12	64.5	17.4	755	T19558	hypothetical prote
13	64.5	17.4	1633	JCS056	polybromo 1 - chic
14	64	17.3	178	T38013	probable integral
15	64	17.3	521	H87111	signal recognition
16	63.5	17.1	100	AH2220	hypothetical prote
17	63.5	17.1	102	B72718	hypothetical prote
18	63.5	17.1	147	B46315	E4 protein - human
19	63.5	17.1	724	F82248	fatty oxidation co
20	63.5	17.1	847	A53800	mixed-lineage prot
21	63	17.0	161	S12246	anther-specific pr
22	63	17.0	265	B45393	polypeptide vi pre
23	63	17.0	273	A46280	prion protein - ch
24	62.5	16.8	295	T28078	hypothetical prote
25	62	16.7	214	B34503	small nuclear ribo
26	62	16.7	231	I53659	Sm-B protein - mou
27	61.5	16.6	197	H86979	hypothetical prote
28	61.5	16.6	211	S55129	transcription fact
29	61.5	16.6	564	AG2823	ABC transporter, m

ALIGNMENTS

RESULT 1

T09067
extensin-like protein NG5 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: T09067
R:Rowen, L.; Mahairas, G.; Qin, S.; Ahearn, M.E.; Dankers, C.; Lasky, S.; Loretz, C.; Sc submitted to the EMBL Data Library, October 1997
A:Description: Sequence of the mouse major histocompatibility locus class III region.
A:Reference number: Z16543
A:Accession: T09067
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-306 <ROW>
A:Cross-references: UNIPROT:O35449; EMBL:AF030001; NID:g2564945; PID:g2564955
C:Genetics:
A:Gene: NG5
A:Map position: 17
A:Introns: 7/1; 186/3; 248/3

Query Match 19.5%; Score 72.5; DB 2; Length 306;

Best Local Similarity 41.2%; Pred. No. 1.1;

Matches 21; Conservative 6; Mismatches 17; Indels 7; Gaps 3;

QY 10 ELSLCFSFFPLLPLHTPVAGRNLFPSLGVPPFLPHPGTTPRARGFL 60

DB 171 QLQPCYAYVP-VYPVGPYAG---GTPGPGVGTSTLPPP---PGPGGLALL 214

RESULT 2

T19888
hypothetical protein C4106.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T19888

R:Cummings, P.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19191

A:Accession: T19888

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-297 <WIL>

A:Cross-references: UNIPROT:O17662; EMBL:Z81047; PIDN:CAB02834.1; GSPDB:GN00023; CESP:CA

C:Experimental source: clone C41G6

C:Genetics:

A:Gene: CESP:C41G6.9

A:Map position: 5

A:Introns: 114/1; 125/2; 168/3; 240/1

C:Superfamily: Caenorhabditis hypothetical protein C49G7.2

Query Match 19.3%;

Best Local Similarity 51.2%; Score 71.5; DB 2; Length 297;

Matches 51.2%; Pred. No. 1.4;

```
Matches 22; Conservative 3; Mismatches 11; Indels 7; Gaps 3;
QY 4 CL-CRGALSICFSPFF---PILLPLHTPVAGRNIGFPESLGVP 42
||| : ||| ||| : ||| : ||| |||
Db 61 CLHCVGAFVDLFFSFAIPALMLPIY---AGVFLGFSRVLGVP 100
||| : ||| ||| : ||| : ||| |||

RESULT 3
T29485
hypothetical protein F08B4.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29485
R:Stelljes, L.; Gattung, S.
submitted to the EMBL Data Library, March 1996
A:Description: The sequence of C. elegans cosmid F08B4.
A:Reference number: Z20625
A:Accession: T29485
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-142 <STE>
A:Cross-references: UNIPROT:P90815; EMBL:U52002; PIDN:AAB37732.1; GSPDB:GN000022; CESP:F08B4
A:Experimental source: strain Bristol N2; clone F08B4
C:Genetics:
A:Gene: CESP:F08B4.7
A:Map position: 4
A:Introns: 39/3; 61/3; 111/3
C:Superfamily: proline-rich protein

Query Match 18.9%; Score 70; DB 2; Length 142;
Best Local Similarity 38.5%; Pred. No. 1;
Matches 15; Conservative 6; Mismatches 16; Indels 2; Gaps 1;
QY 19 PLLAPLHTPVAGRNIGFPESLGVPFPLPHPGGTPRAPGL 57
||| : ||| : ||| : ||| : ||| : |||
Db 92 PWMAP--RPFPGVGFFGAPGMPFFPGPGMGAGPFCM 128
||| : ||| : ||| : ||| : ||| : |||

RESULT 4
JC5319
macrolide-lincosamide-streptogramin B resistance determinant - Streptomyces fradiae
A:Alternate names: Tlrd protein
C:Species: Streptomyces fradiae
C>Date: 01-May-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jul-2004
C:Accession: JC5319
R:Gandecha, A.R.; Cundliffe, E.
Gene 180, 173-176, 1996
A:Title: Molecular analysis of tlrd, an MLS resistance determinant from the tylosin prod
A:Reference number: JC5319; MUID:97128823; PMID:8973363
A:Accession: JC5319
A:Molecule type: DNA
A:Residues: 1-327 <GAN>
A:Cross-references: UNIPROT:P97178; EMBL:X97721; NID:g1781317; PIDN:CAA66307.1; PID:g1781317
C:Comment: This protein monomethylates residue A-2058 in 23S rRNA generating N6-methylad
C:Genetics:
A:Gene: tlrd
C:Superfamily: rRNA (adenine-N6-)-methyltransferase

Query Match 18.7%; Score 69.5; DB 2; Length 327;
Best Local Similarity 45.0%; Pred. No. 2.6;
Matches 18; Conservative 2; Mismatches 17; Indels 3; Gaps 1;
QY 19 PLLAPLHTPVAGRNIGFPESLGVP---PFLPHPGGTPRAP 55
||| : ||| : ||| : ||| : ||| : |||
Db 276 PRVAPRRPPGASRRSGAESRGVPRRWGVWPAPGGRGAP 315
||| : ||| : ||| : ||| : ||| : |||

RESULT 5
B70908
hypothetical protein Rv0591 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: B70908
```

```
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70908
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-481 <COL>
A:Cross-references: UNIPROT:O07787; GB:Z97182; GB:AL123456; NID:g3250720; PIDN:CAB09961.1
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv0591

Query Match 18.1%; Score 67; DB 2; Length 481;
Best Local Similarity 40.5%; Pred. No. 7.3;
Matches 15; Conservative 3; Mismatches 13; Indels 6; Gaps 1;
QY 27 PVAGRNIG-----FPESLGVPFPLPHPGGTPRAPGL 57
||| : ||| : ||| : ||| : ||| : |||
Db 400 PCAGATVGPFGDFPAPLDVQSPNPDPGPPPTPGI 436
||| : ||| : ||| : ||| : ||| : |||

RESULT 6
Q08E44
BDLF2 protein - human herpesvirus 4 (strain B95-8)
C:Species: human herpesvirus 4, Epstein-Barr virus
C>Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004
C:Accession: G43044; A03788; S33041
R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
Mol. Biol. Med. 1, 21-45, 1983
A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
A:Reference number: A93065; MUID:85035713; PMID:6052825
A:Accession: G43044
A:Molecule type: DNA
A:Residues: 1-420 <BAN>
A:Cross-references: UNIPROT:P03225; EMBL:V01555; NID:G59074; PIDN:CAA24836.1; PID:g13349
R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H.
Nature 310, 207-211, 1984
A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
A:Reference number: A03794; MUID:84270667; PMID:6087149
A:Contents: annotation; protein coding region
C:Superfamily: human herpesvirus 4 BDLF2 protein

Query Match 17.9%; Score 66.5; DB 1; Length 420;
Best Local Similarity 43.9%; Pred. No. 7.3;
Matches 18; Conservative 1; Mismatches 13; Indels 9; Gaps 2;
QY 13 LCFSPFPLLLPLHTPVA-----GRNLGFPESLGVP--PF 44
||| : ||| : ||| : ||| : ||| : |||
Db 204 LAFGNPLFLPSFMPVCAKVLKGRDFGVPLSYGCTPNPF 244
||| : ||| : ||| : ||| : ||| : |||

RESULT 7
D70747
probable ffh protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: D70747
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70747
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-525 <COL>
A:Cross-references: UNIPROT:Q10963; GB:Z74697; GB:AL123456; NID:g3261602; PIDN:CAA98978.
```

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: ffh

C:Superfamily: signal recognition particle 54K protein

Query Match 17.8%; Score 66; DB 2; Length 525;

Best Local Similarity 45.2%; Pred. No. 10;

Matches 19; Conservative 3; Mismatches 16; Indels 4; Gaps 3;

QY 19 PLL--PLHTPVAGNMG-FPESLGVPPFLPHPGGTPRA-PG 56

DB 131 PLLVACDLQRAVNVQLVVGKGVVFAHPHGASPSGPG 172

RESULT 8

S68470

Ca2+/calmodulin-dependent protein kinase (BC 2.7.1.123) II beta-3 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: S68470; I53501

R:Urquidí, V.; Ashcroft, S.J.H.

FBBS Lett. 358, 23-26, 1995

A:Title: A novel pancreatic beta-cell isoform of calcium/calmodulin-dependent protein kinase II

A:Reference number: I53501; MUID:95121451; PMID:7821422

A:Accession: S68470

A:Molecule type: mRNA

A:Residues: 1-589 <URQ>

A:CROSS-references: UNIPROT:O63094; EMBL:X83375; NID:g603580; PIDN:CAA58289.1; PID:g6035

A:Experimental source: pancreatic islets

C:Complex: heteromultimer composed of 10-12 chains (alpha, beta, gamma, delta) generally

C:Superfamily: Ca2+/calmodulin-dependent protein kinase II; protein kinase homology

C:Keywords: alternative splicing; ATP; autophosphorylation; calmodulin binding; phosphop

F:12-272/Domain: protein kinase homology <KIN>

F:20-28/Region: protein kinase ATP-binding motif

F:287-311/Region: calmodulin binding #status predicted

F:43.61.136.138/Active site: Lys, Glu, Asp, Lys #status predicted

F:287,306/Binding site: phosphate (Thr) (covalent) (by autophosphorylation) #status pred

Query Match 17.8%; Score 66; DB 2; Length 589;

Best Local Similarity 33.9%; Pred. No. 12;

Matches 21; Conservative 3; Mismatches 18; Indels 20; Gaps 2;

QY 4 CLCGAELSLCFSPFLLLPLHTP-----VAGNMGFPESLGVPPFLPHPGGTPRA 54

DB 400 CLSPG-----LLGLPTSPRISDILNSVRGSGTPEARGLPVGPVPPCPSPPL 448

QY 55 PG 56

DB 449 PG 450

RESULT 9

D70815

probable fadB protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: D70815

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: D70815

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-720 <COL>

A:CROSS-references: UNIPROT:O53872; GB:AL022004; GB:AL123456; NID:g3261550; PIDN:CAA1768

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: fadB

C:Superfamily: enoyl-CoA hydratase/3-hydroxyacyl-CoA dehydrogenase; 3-hydroxyacyl-CoA de

F:16-187/Domain: enoyl-CoA hydratase homology <ECH>

F:325-608/Domain: 3-hydroxyacyl-CoA dehydrogenase homology <HCD>

Query Match 17.8%; Score 66; DB 2; Length 720;

Best Local Similarity 45.5%; Pred. No. 14;

Matches 15; Conservative 4; Mismatches 10; Indels 4; Gaps 1;

QY 36 PESLGVPPF----LPHPGGTPRAPGLFLLFSPF 64

DB 213 PDGAGVQWDKKGKYMKGCTPSSPGLAAILPSF 245

RESULT 10

E75447

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: E75447

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: E75447

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-257 <WHI>

A:CROSS-references: UNIPROT:Q9RVL4; GB:AE001953; GB:AE000513; NID:g6458740; PIDN:AAF1059

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1013

A:Map position: 1

Query Match 17.7%; Score 65.5; DB 2; Length 257;

Best Local Similarity 42.9%; Pred. No. 5.8;

Matches 18; Conservative 1; Mismatches 18; Indels 5; Gaps 1;

QY 17 FFPDLLPLHTPVAGNMGFPESLGVPPFLPHPGGTPRAPGLF 58

DB 109 FFPDLLVLTLGLSLVAGIPHGYG-----EFRGGTGLGGLF 145

RESULT 11

E72501

hypothetical protein APE1987 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C:Accession: E72501

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: E72501

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-102 <KAW>

A:CROSS-references: UNIPROT:Q9YAF2; DBJ:AP000063; NID:g5105654; PIDN:BAA80997.1; PID:d

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1987

Query Match 17.4%; Score 64.5; DB 2; Length 102;

Best Local Similarity 43.2%; Pred. No. 3;

Matches 16; Conservative 4; Mismatches 14; Indels 3; Gaps 2;

QY 18 FPLLPL--LHTPVAGNMGFPESLGVPPFLPHPGGTP 52

DB 29 YPLHAPYLLLYKPGSGRGYGAPPPI-VPPGYPHPNLLP 64

RESULT 12

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2005, 20:36:18 ; Search time 165 Seconds
(without alignments)
157.048 Million cell updates/sec

Title: US-10-649-857-45

Perfect score: 371
Sequence: 1 MGKLCRGAEISLCSFPPL.....PGGTRAPGLFLLLFSFWAV 67

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_18Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	371	100.0	67	2 AAW83944	AAW83944 Human sec
2	371	100.0	150	2 AAY60549	Aay60549 Human nor
3	282	76.0	51	2 AAY48336	Aay48336 Human pro
4	76	20.5	126	4 AAO00972	Aao00972 Human pol
5	74.5	20.1	200	7 ADD69583	Add69583 Human REM
6	74.5	20.1	347	8 ADQ65482	Adq65482 Novel hum
7	73	19.7	200	8 ADR10217	Adr10217 Human pro
8	72	19.4	223	4 AAO06156	Aao06156 Human pol
9	71	19.1	135	4 AAO01562	Aao01562 Human pol
10	71	19.1	672	4 AAU28230	Aau28230 Novel hum
11	71	19.1	672	4 AAU28232	Aau28232 Novel hum
12	71	19.1	672	4 AAU28231	Aau28231 Novel hum
13	70.5	19.0	142	4 AAO00501	Aao00501 Human pol
14	69	18.6	107	4 ABG12930	Abg12930 Novel hum
15	67	18.1	100	5 ABK34374	Abk34374 Novel hum
16	67	18.1	117	4 AAO13083	Aao13083 Human pol
17	67	18.1	481	5 ABJ04679	Abj04679 Mycobacte
18	66.5	17.9	142	7 ADB64520	Adb64520 Human pro
19	66.5	17.9	461	5 ADK34685	Adk34685 Novel hum
20	66.5	17.9	806	6 ABO07125	Ab07125 Novel hum
21	66	17.8	131	4 AAO04859	Aao04859 Human pol
22	66	17.8	179	4 ABG18182	Abg18182 Novel hum
23	66	17.8	517	6 ABU34062	Abu34062 Protein e
24	66	17.8	525	6 ABU34661	Abu34661 Protein e
25	66	17.8	525	6 ABU36878	Abu36878 Protein e

26	66	17.8	712	6 ABU34588	Abu34588 Protein e
27	66	17.8	720	6 ABU36538	Abu36538 Protein e
28	65.5	17.7	135	4 AAO01708	Aao01708 Human pol
29	65	17.5	421	7 ADI21224	Adi21224 Novel hum
30	65	17.5	422	4 ABG22990	Abg22990 Novel hum
31	65	17.5	452	8 ADQ66221	Adq66221 Novel hum
32	64.5	17.4	130	4 AAO01554	Aao01554 Human pol
33	64.5	17.4	424	8 ADG22716	Adg22716 Cyanophag
34	64.5	17.4	426	7 ABO65478	Ab065478 Klebsiell
35	64.5	17.4	692	6 ABU28111	Abu28111 Protein e
36	64	17.3	106	4 AAO07378	Aao07378 Human pol
37	64	17.3	320	5 AAU78461	Aau78461 Mouse bet
38	64	17.3	413	7 ADJ83115	Adj83115 Ovine con
39	64	17.3	521	6 ABU35972	Abu35972 Protein e
40	64	17.3	1494	5 AAU78460	Aau78460 Mouse bet
41	63.5	17.1	123	4 AAO04222	Aao04222 Human pol
42	63.5	17.1	336	7 ADB63875	Adb63875 Human pro
43	63.5	17.1	710	6 ABU41050	Abu41050 Protein e
44	63.5	17.1	724	6 ABU49290	Abu49290 Protein e
45	63.5	17.1	751	7 ADF03896	Adf03896 Bacterial

ALIGNMENTS

RESULT 1
AAW83944
ID AAW83944 standard; protein; 67 AA.
XX
AC AAW83944;
XX
DT 28-JAN-1999 (first entry)
XX
DE Human secreted protein from gene 14 clone HNGBV36.
XX

Secreted protein; gene therapy; protein therapy; diagnosis; treatment;
central nervous system; CNS; immune system; cancer; trauma; liver;
reproductive disorder; congenital malformation; degenerative disease;
inflammatory disease; neoplasia; metabolic disorder; testis; placenta;
brain; T cell; spleen; lung; heart; rhabdomyosarcoma; endocrine system;
endocrinopathy; endocrine polyglandular syndrome; endocrinoma; sepsis;
endocrine ophthalmopathy; osteoclastoma; bacterial infection; bone.

Homo sapiens.

Key Location/Qualifiers
FT Peptide 1..32
FT Protein /note= "signal peptide"
FT Protein /note= "mature secreted protein"

WO9845712-A2.

15-OCT-1998.

07-APR-1998; 98WO-US006801.

08-APR-1997; 97US-0042726P.

08-APR-1997; 97US-0042727P.

08-APR-1997; 97US-0042728P.

08-APR-1997; 97US-0042754P.

08-APR-1997; 97US-0042825P.

30-MAY-1997; 97US-0048058P.

30-MAY-1997; 97US-0048070P.

30-MAY-1997; 97US-0048184P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Yu G, Ni J, Feng P;

WPI; 1998-594496/50.

DR N-PSDB; AAV69624.

XX

PT New isolated human genes and secreted polypeptide(s) they encode - useful
 PT for the diagnosis and treatment of e.g. cancers, CNS disorders, immune
 PT system disorders, inflammatory disease and bacterial infections.

XX Claim 11; Page 123; 142pp; English.

PS This sequence represents a human secreted protein encoded by a nucleic
 CC acid molecule designated Gene 14 from the human cDNA clone HNG8V36
 CC (deposited as clone ATCC 97955 and ATCC 209074). The gene is expressed
 CC primarily in breast cancer, pituitary and activated T cells and to a
 CC lesser extent in frontal cortex and breast and is useful for diagnosis
 CC and treatment of breast cancer and growth disorders. The invention
 CC relates to 20 novel genes and their fragments (AAV69611 to AAV69630) and
 CC corresponding secreted proteins (AAW83931 to AAW83950) which are useful
 CC for preventing, treating or ameliorating medical conditions e.g. by
 CC protein of gene therapy. Also pathological conditions can be diagnosed by
 CC determining the amount of the new polypeptides in a sample or by
 CC determining the presence of mutations in the polynucleotides. Specific
 CC uses are based on which tissues they are most highly expressed in, and
 CC include developing products for the diagnosis or treatment of central
 CC nervous system (CNS) and immune system diseases, reproductive disorders,
 CC cancers, congenital malformations, degenerative diseases, trauma,
 CC inflammatory disease, neoplasia, metabolic disorders, diseases in testes,
 CC placenta, liver, brain and activated T cells, spleen diseases, lung
 CC diseases, heart diseases, rhabdomyosarcoma and disorders of the endocrine
 CC system or other endocrinopathies, e.g. endocrine polyglandular syndrome,
 CC endocrinoma, and endocrine ophthalmopathy, osteoclastoma and other bone
 CC remodelling disorders, bacterial infections and sepsis. The polypeptides
 CC are also useful for identifying their binding partners

XX Sequence 67 AA;

Query Match 100.0%; Score 371; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 1.8e-34;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKCLCRGAELSLCFSPFLLPLHTPVAGRNIGFPESLGVPPFLPHPGGTPRAGLFL 60
 DB 1 MGKCLCRGAELSLCFSPFLLPLHTPVAGRNIGFPESLGVPPFLPHPGGTPRAGLFL 60
 QY 61 LFSFWAV 67
 DB 61 LFSFWAV 67

RESULT 2
 AAY60549
 ID AAY60549 standard; protein; 150 AA.

XX AC AAY60549;

XX DT 31-JAN-2000 (first entry)

XX Human normal bladder tissue EST encoded protein 221.

XX Human; bladder; treatment; EST; expressed sequence tag; cytostatic;
 XX cancer; gene therapy.

XX OS Homo sapiens.

XX PN DE19818620-A1.

XX PD 28-OCT-1999.

XX PF 21-APR-1998; 98DE-01018620.

XX PR 21-APR-1998; 98DE-01018620.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 XX WPI; 1999-602416/52.

DR N-PSDB; AAZ42232.

XX New polypeptides and their nucleic acids, useful for treatment of bladder
 PT tumor and identification of therapeutic agents.

PS Claim 23; Page 334; 366pp; German.

XX This invention describes novel polypeptide fragment sequences (I) and
 CC their encoding nucleic acids (II) which are highly expressed in normal
 CC bladder tissue and have cytostatic activity. (II) are used for
 CC recombinant expression of (I) and to isolate complete genes. (I) are used
 CC to identify agents suitable for the treatment of bladder tumours, to
 CC directly treat this form of cancer (including expression from gene
 CC therapy vectors), or are used in a preparation for cancer treatment. (I)
 CC is also used for the generation of specific antibodies. (II) are
 CC identified by assembling ESTs (expressed sequence tags) from a particular
 CC tissue type before comparison of expression patterns. This allows a
 CC significantly longer fragment of the gene to be revealed, and therefore
 CC reduces the number of failures because of ESTs from different libraries
 CC representing different parts of the same unknown gene distorting the
 CC estimated frequency of occurrence in a particular tissue. AAY60329-Y60591
 CC represent protein fragments encoded by the human normal bladder tissue
 CC cDNA library derived EST fragments represented in AAZ42122-Z42248

XX Sequence 150 AA;

Query Match 100.0%; Score 371; DB 2; Length 150;
 Best Local Similarity 100.0%; Pred. No. 4e-34;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGKCLCRGAELSLCFSPFLLPLHTPVAGRNIGFPESLGVPPFLPHPGGTPRAGLFL 60
 DB 84 MGKCLCRGAELSLCFSPFLLPLHTPVAGRNIGFPESLGVPPFLPHPGGTPRAGLFL 143
 QY 61 LFSFWAV 67
 DB 144 LFSFWAV 150

RESULT 3

AAY48336

ID AAY48336 standard; protein; 51 AA.

XX AC AAY48336;

XX DT 08-DEC-1999 (first entry)

XX Human prostate cancer-associated protein 33.

XX Expressed sequence tag; EST; prostate; tumor; treatment; gene therapy;
 XX cancer; tissue specificity; human.

XX OS Homo sapiens.

XX PN DE19811194-A1.

XX PD 16-SEP-1999.

XX PF 10-MAR-1998; 98DE-01011194.

XX PR 10-MAR-1998; 98DE-01011194.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

XX Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
 XX WPI; 1999-519629/44.
 XX N-PSDB; AAZ33494.

XX New nucleic acid expressed at high level in normal prostatic tissue and
 PT encoded polypeptides, used to treat cancer and screen for therapeutic
 PT agents.

PS Claim 22; 137; 194pp; German.

XX This invention describes novel nucleic acid sequences (A) that are expressed at high level in normal prostatic tissue. Polypeptides (I) encoded by (A) are used; (a) for identifying agents for treatment of prostatic cancer and (b) for therapy of prostate cancer, optionally where expressed by gene therapy methods. (A) is also used to isolate full-length genes (for gene therapy) and for recombinant production of (I), which can be used to raise specific antibodies. (A) are identified by assembly of ESTs (expressed sequence tags) before these are analyzed for expression pattern (tissue specificity). This approach eliminates many of the false results, as regards tissue specificity, associated with known methods that use single (usually short) ESTs. AA148304-Y48456 represent peptides encoded by the expressed sequence tags described in the method of the invention

XX Sequence 51 AA;

Query Match 76.0%; Score 282; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.7e-24;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 FPLLPLHTPVAGNLCGPPSLGVPPPLPHPGGTPRAGLFLLSFWAV 67
|||||
Db 1 FPLLPLHTPVAGNLCGPPSLGVPPPLPHPGGTPRAGLFLLSFWAV 51

RESULT 4

AAO00972

ID AAO00972 standard; protein; 126 AA.

XX AAO00972;

XX 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 14864.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

XX 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSBQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI: 2001-514838/56.

XX N-PSDB; AA180903.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 14864; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 126 AA;

Query Match 20.5%; Score 76; DB 4; Length 126;
Best Local Similarity 37.9%; Pred. No. 1.1;
Matches 22; Conservative 3; Mismatches 21; Indels 12; Gaps 3;

Qy 18 FPLLPLHTPVAGNLCGPPSLGV-PPFLPH-----PGGTPRAGLFLLSF 64
|||||
Db 8 FP-LSPLYTPPHGHLGPPOTFGAGPPAHKSHQVGRQKRGFGVPPRPAPFLFFFF 64

RESULT 5

ADD69583

ID ADD69583 standard; protein; 200 AA.

XX ADD69583;

XX 15-JAN-2004 (first entry)

XX Human REMAP protein - SEQ ID 12.

XX human; receptor and membrane-associated protein; REMAP; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic; neuroprotective; cerebroprotective; anti-HIV; antiallergic; antiinflammatory; thymimetic; cell proliferative; cancer; atherosclerosis; neurological; epilepsy; Huntington's disease; stroke; immune; inflammatory; AIDS; allergy; developmental; hypothyroidism; Cushing's syndrome; infection.

XX Homo sapiens.

XX WO2003048305-A2.

XX 12-JUN-2003.

XX 13-NOV-2002; 2002WO-US036759.

XX 13-NOV-2001; 2001US-0333097P.

XX 15-NOV-2001; 2001US-0335274P.

XX 14-DEC-2001; 2001US-0340542P.

XX 18-DEC-2001; 2001US-0342166P.

XX 11-JAN-2002; 2002US-0347580P.

XX 14-JAN-2002; 2002US-0348687P.

XX (INCY-) INCYTE GENOMICS INC.

XX Duggan BM, Yang J, Gietzen KJ, Lee SY, Tang YT, Azimzai Y; Chawla NK, Warren BA, Barroso I, Becha SD, Yue H, Lehr-Mason PM; Thangavelu K, Lee S, Emerling BM, Kable AE, Khare R, Baughn MR; Gandhi AR, Tran UK, Richardson TW, Marquis JP, Lal PG, Forsythe IJ; Lee EA, Swarnakar A, Kallick DA, Griffin JA, Elliott VS, Gorvad AE; Hafalia AJA, Ison CH, Jin P, Jiang X, Jackson AA, Bhatia U; Burrill JD, Blake JJ, Ho A, Zheng W, Gao J;

XX WPI: 2003-513744/48.

XX N-PSDB; ADD69630.

XX New human receptors and membrane-associated proteins (REMAP), useful for diagnosing, treating and preventing diseases or conditions associated with the aberrant REMAP expression e.g. cancer, AIDS, atherosclerosis, or infections.

XX Claim 1; SEQ ID NO 12; 298pp; English.

XX The invention relates to a novel isolated polypeptide comprising a human receptor and membrane-associated protein (REMAP). The polypeptide of the invention demonstrates cytostatic, antiarteriosclerotic, anticonvulsant, nootropic, neuroprotective, cerebroprotective, anti-HIV, antiallergic,

QY 17 FPLLPLHT--PVAGRNIGFESLGVPPFLPHPGGTPR-APGLFLLLSFW 65
 Db 130 FSPCTIPVHSGFPVEG---GAPGQSVFPHLTHTLGDPRLPLGALTDIAEW 178

RESULT 8
 AAO06156
 ID AAO06156 standard; protein; 223 AA.
 AC AAO06156;
 XX
 XX 06-NOV-2001 (first entry)
 XX
 XX Human polypeptide SEQ ID NO 20048.
 DE
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US004927.
 XX
 PR 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX
 PA (HYSE-) HYSSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-514838/56.
 DR N-PSDB; AAI86087.
 XX
 XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.
 PT
 XX Claim 20; SEQ ID NO 20048; 1399pp + Sequence Listing; English.
 XX
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 223 AA;
 Query Match 19.4%; Score 72; DB 4; Length 223;
 Best Local Similarity 45.5%; Pred. No. 5.4;
 Matches 15; Conservative 0; Mismatches 10; Indels 8; Gaps 1;

QY 23 PLHTPVAGNIGFESLGVPPFLPHPGGTPRAP 55
 Db 121 PSHVP-----PHPLAFPPSLPHPGPASRAP 145

RESULT 9
 AAO01562
 ID AAO01562 standard; protein; 135 AA.
 XX
 AC AAO01562;

XX 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 15454.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US004927.
 XX
 PR 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX
 PA (HYSE-) HYSSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-514838/56.
 DR N-PSDB; AAI81493.
 XX
 XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.
 PT
 XX Claim 20; SEQ ID NO 15454; 1399pp + Sequence Listing; English.
 XX
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 135 AA;
 Query Match 19.1%; Score 71; DB 4; Length 135;
 Best Local Similarity 36.4%; Pred. No. 4.2;
 Matches 24; Conservative 4; Mismatches 22; Indels 16; Gaps 5;

QY 7 RGABLSLCFSF-----PFLPLPLHTPVAGNIGFPE--SLGVPPFL-PHP-----GGTP 52
 Db 13 KGIDGFLFFFGKFKFFLAPGAP--GGLFPEASSPGLNFFFWNPPEKKKGPP 70

QY 53 RAPGLF 58
 Db 71 PPGGFF 76

RESULT 10
 AAU28230
 ID AAU28230 standard; protein; 672 AA.
 XX
 AC AAU28230;
 XX
 XX 18-DEC-2001 (first entry)
 XX
 DE Novel human secretory protein, Seq ID No 587.
 XX
 KW Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;

KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.
 XX
 OS Homo sapiens.
 XX
 PN WO200166689-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 05-MAR-2001; 2001WO-US004942.
 XX
 PR 07-MAR-2000; 2000US-00519705.
 PR 19-MAY-2000; 2000US-00574454.
 PR 17-JUN-2000; 2000US-00596193.
 PR 14-JUL-2000; 2000US-00616847.
 PR 19-SEP-2000; 2000US-00665363.
 PR 20-OCT-2000; 2000US-00693267.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;
 XX WPI; 2001-589934/66.
 DR N-PSDB; AAS45130.
 XX
 PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders.
 XX
 PS Example 2; SEQ ID NO 587; 107pp; English.
 XX
 CC The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention
 XX
 SQ Sequence 672 AA;

Query Match 19.1%; Score 71; DB 4; Length 672;
 Best Local Similarity 40.7%; Pred. No. 21;
 Matches 22; Conservative 1; Mismatches 17; Indels 14; Gaps 3;
 QY 19 PLLPLHTPVAGRNLCF--PESIG-----VPPPLPHPGGT----PRAPGLP 58
 Db 589 PLQPFPHPPHGLACGFLLPSSVGGRLDCGFCFLGAQGGAWFHPRPTAP 642

RESULT 11

AAU28232
 ID AAU28232 standard; protein; 672 AA.

XX AAU28232;

DT 18-DEC-2001 (first entry)

XX Novel human secretory protein, Seq ID No 589.

XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.

OS Homo sapiens.

PN WO200166689-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001WO-US004942.

XX 07-MAR-2000; 2000US-00519705.

PR 19-MAY-2000; 2000US-00574454.

PR 17-JUN-2000; 2000US-00596193.

PR 14-JUL-2000; 2000US-00616847.

PR 19-SEP-2000; 2000US-00665363.

PR 20-OCT-2000; 2000US-00693267.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren F, Ma Y, Zhou P;
 PI Zhao QA, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AJ, Wang J;

XX WPI; 2001-589934/66.

DR N-PSDB; AAS45132.

XX Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis and treatment of
 PT cancer, neurological, inflammatory, and autoimmune disorders.
 XX
 PS Example 2; SEQ ID NO 589; 107pp; English.

XX The invention relates to novel isolated human secreted polypeptides (I)
 CC and polynucleotides (II). (I) and (II) are useful for treating
 CC inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 CC ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 CC involved in increasing haematopoiesis, stem cell survival, bone growth
 CC and remodeling. (I), (II) and modulators of (II) are useful for
 CC prophylaxis or treatment of one or more cancers. (II) is also useful for
 CC creating transgenic animals useful for studying the in vivo activities of
 CC the polypeptide as well as for studying modulators of the polypeptides.
 CC (I) induces the proliferation of neural cells and regeneration of nerve
 CC and brain tissue and is useful for the treatment of central and
 CC peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 CC Parkinson's disease, Huntington's disease, and amyotrophic lateral
 CC sclerosis. In addition, (I) is involved in chemotactic or chemokinetic

CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth, and in tissue repair, healing of burns, incisions, ulcers,
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders,
 CC or periodontal disease. Furthermore, (I) is also useful for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, various immune deficiencies and
 CC disorders including severe combined immunodeficiency (SCID), bacterial or
 CC fungal infections, autoimmune disorders e.g. multiple sclerosis,
 CC rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 CC reactions and conditions, such as asthma or other respiratory problems.
 CC In addition, (I) affects biorhythms or circadian cycles of rhythms,
 CC fertility, metabolism, catabolism, anabolism, storage or elimination of
 CC dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 CC analgesic effects or other pain reducing effects, immunoglobulin like
 CC activity and can act as an antigen in a vaccine composition to raise an
 CC immune response. AAU28020-AAU28395 represent novel human secreted protein
 CC amino acid sequences of the invention
 XX
 SQ Sequence 672 AA;

Query Match 19.1%; Score 71; DB 4; Length 672;
 Best Local Similarity 40.7%; Pred. No. 21;
 Matches 22; Conservative 1; Mismatches 17; Indels 14; Gaps 3;
 Qy 19 PLLPLHTPVAGRNIGF--PESLG-----VPPFLPHPGGT----PRAPGLF 58
 Db 589 PLQLPFHPPLGLACGFLPSSVGGRLCGFVPCFLGAGGAWFHPRPTAF 642

RESULT 12
 AAU28231
 ID AAU28231 standard; protein; 672 AA.

AC AAU28231;
 XX
 DT 18-DEC-2001 (first entry)
 XX

DE Novel human secretory protein, Seq ID No 588.

XX Human; secreted protein; arthritis; Crohn's disease; sepsis; shock;
 KW ischaemia-reperfusion injury; haematopoiesis; cancer; neuropathy;
 KW transgenic animal; Alzheimer's disease; Parkinson's disease; burn;
 KW amyotrophic lateral sclerosis; platelet disorder; thrombocytopenia;
 KW ulcer; osteoporosis; bone degenerative disorder; periodontal disease;
 KW gut protection; lung; liver fibrosis; immune deficiency; infection;
 KW severe combined immunodeficiency; SCID; autoimmune disorder; allergy;
 KW multiple sclerosis; rheumatoid arthritis; diabetes mellitus; asthma;
 KW fertility; analgesic; pain; antigen.

XX Homo sapiens.

OS

FN WO200016689-A2.

XX 13-SEP-2001.

XX 05-MAR-2001; 2001WO-US004942.

XX 07-MAR-2000; 2000US-00519705.

PR 19-MAY-2000; 2000US-00574454.

PR 17-JUN-2000; 2000US-00596193.

PR 14-JUL-2000; 2000US-00616847.

PR 19-SEP-2000; 2000US-00665363.

PR 20-OCT-2000; 2000US-00693267.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Xu C, Wehrman T, Ren P, Ma Y, Zhou P;

PI Zhao QH, Yang Y, Drmanac RT, Zhang J, Chen R, Xue AU, Wang J;

XX WPI; 2001-589934/66.

DR N-PSDB; AAS45131.

XX

Novel polypeptides and nucleic acids obtained from cDNA libraries
 prepared from various human tissues, for diagnosis and treatment of
 cancer, neurological, inflammatory, and autoimmune disorders.

Example 2; SEQ ID NO 588; 107pp; English.

The invention relates to novel isolated human secreted polypeptides (I)
 and polynucleotides (II). (I) and (II) are useful for treating
 inflammatory conditions such as arthritis, nephritis, Crohn's disease,
 ischaemia-reperfusion injury, shock, sepsis, immune responses, and is
 involved in increasing haematopoiesis, stem cell survival, bone growth
 and remodeling. (I), (II) and modulators of (II) are useful for
 prophylaxis or treatment of one or more cancers. (II) is also useful for
 creating transgenic animals useful for studying the in vivo activities of
 the polypeptide as well as for studying modulators of the polypeptides.
 (I) induces the proliferation of neural cells and regeneration of nerve
 and brain tissue and is useful for the treatment of central and
 peripheral nervous system diseases and neuropathies, such as Alzheimer's,
 Parkinson's disease, Huntington's disease, and amyotrophic lateral
 sclerosis. In addition, (I) is involved in chemotactic or chemokinetic
 activity, regulation of haematopoiesis and is useful for treating myeloid
 or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 tissue growth, and in tissue repair, healing of burns, incisions, ulcers,
 for treating osteoporosis, osteoarthritis, bone degenerative disorders,
 or periodontal disease. Furthermore, (I) is also useful for gut
 protection or regeneration and treatment of lung or liver fibrosis,
 reperfusion injury in various tissues, various immune deficiencies and
 disorders including severe combined immunodeficiency (SCID), bacterial or
 fungal infections, autoimmune disorders e.g. multiple sclerosis,
 rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic
 reactions and conditions, such as asthma or other respiratory problems.
 In addition, (I) affects biorhythms or circadian cycles of rhythms,
 fertility, metabolism, catabolism, anabolism, storage or elimination of
 dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides
 analgesic effects or other pain reducing effects, immunoglobulin like
 activity and can act as an antigen in a vaccine composition to raise an
 immune response. AAU28020-AAU28395 represent novel human secreted protein
 amino acid sequences of the invention

XX Sequence 672 AA;

Query Match 19.1%; Score 71; DB 4; Length 672;
 Best Local Similarity 40.7%; Pred. No. 21;
 Matches 22; Conservative 1; Mismatches 17; Indels 14; Gaps 3;

Qy 19 PLLPLHTPVAGRNIGF--PESLG-----VPPFLPHPGGT----PRAPGLF 58
 Db 589 PLQLPFHPPLGLACGFLPSSVGGRLCGFVPCFLGAGGAWFHPRPTAF 642

RESULT 13

AAO00501

ID AAO00501 standard; protein; 142 AA.

XX AC AAO00501;

XX 06-NOV-2001 (first entry)

DT Human polypeptide SEQ ID NO 14393.

DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX FN WO200164835-A2.

XX 07-SEP-2001.

XX PD


```

PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
DR N-PSDB; AAI80432.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 20; SEQ ID NO 14393; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 142 AA;
SQ
Query Match 19.0%; Score 70.5; DB 4; Length 142;
Best Local Similarity 37.3%; Pred. No. 5.1;
Matches 19; Conservative 1; Mismatches 20; Indels 11; Gaps 2;
QY 19 PLLALPLH-----TPVAGNRLGFPSLGVPPF-----LPHFGGTPRAPGLF 58
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 29 FLSPPLMGSKVPVXPKVGRPLPGXAGKPPFFLKFQPPFPFGGPRCPPLF 79
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 14
ABG12930
ID ABG12930 standard; protein; 107 AA.
XX
AC ABG12930;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #12921.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS77117.
XX

```

```

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 43289; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG10377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 107 AA;
SQ
Query Match 18.6%; Score 69; DB 4; Length 107;
Best Local Similarity 36.2%; Pred. No. 5.6;
Matches 25; Conservative 5; Mismatches 17; Indels 22; Gaps 5;
QY 8 GAELSLC--FSFPPL-----LLPLHTPVAGNRLGFPSLGVPPPLPHFGGTPRAPGL 57
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 GLRLSLAPPFCMPFLPTTXXHSVLPASFPAA---LGIPQARAVTPPV-----SPXL 50
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 58 FLLLFQFWA 66
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 51 -LASLSFWA 58
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 15
ADK34374
ID ADK34374 standard; protein; 100 AA.
XX
AC ADK34374;
XX
DT 06-MAY-2004 (first entry)
XX
DE Novel human polypeptide SeqID6456.
XX
KW antiarthritic; antiparkinsonian; neuroprotective; nootropic;
KW immunosuppressive; cytostatic; antipsoriatic; antiinflammatory;
KW antibacterial; antiviral; antifungal; antiparasitic; gene therapy;
KW arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;
KW psoriasis; inflammatory bowel disease; infection; bacteria; virus;
KW fungus; parasite; human.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 1. .100
XX /label= OTHER
XX /note= "OTHER= All Xaa's in this sequence are unknown
XX amino acids or the site of a stop codon within the DNA
XX sequence"
XX
XX WO200216439-A2.
XX
XX 28-FEB-2002.
XX

```


XX 05-MAR-2001; 2001WO-US004941.
XX
XX 07-MAR-2000; 2000US-00519705.
PR 19-MAY-2000; 2000US-00574454.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2002-280918/32.
XX
XX Isolated polynucleotide encoding bone marrow derived polypeptides useful
PT for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's
PT disease, and inflammatory bowel disease.
XX
XX Claim 20; SEQ ID NO 6456; 504pp; English.
XX
XX This invention relates to a novel isolated polynucleotide comprising a
CC nucleotide sequence selected from one of 1680 sequences, a mature protein
CC coding portion of them, an active domain of them and their complementary
CC sequences. The invention may be useful for the production of compounds
CC with an antiarthritic, antiparkinsonian, neuroprotective, neurotropic,
CC immunosuppressive, cytostatic, anticancer, antiparasitic, antiinflammatory,
CC antibacterial, antiviral, antifungal or antiparasitic activity. In
CC addition, the disclosed sequences may be useful for gene therapy. The
CC polypeptides or their antibodies are useful for treating many diseases
CC such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,
CC psoriasis, inflammatory bowel disease and infections caused by bacteria,
CC viruses, fungi or parasites. The present sequence is that of a human
CC polypeptide of the invention.
XX
XX Sequence 100 AA;
SQ
Query Match 18.1%; Score 67; DB 5; Length 100;
Best Local Similarity 40.4%; Pred. No. 8.9;
Matches 21; Conservative 1; Mismatches 16; Indels 14; Gaps 3;
QY 17 FFFLLPLHTPVAGRNIGPESLGVF-----PFLPHPGG--TPRAPGLFLL 60
DB 18 FFPQAPNP-----LGPFGGLGPNPPPPGFFPPFGGKPPPNKGFFLL 63

Search completed: August 29, 2005, 20:52:23
Job time : 168 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)